Illumination-robust Recognition and Inspection in Carbide Insert Production

R. Schmitt[1] and Y. Cai[1]
[1] Laboratory for Machine Tools and Production Engineering, RWTH Aachen University, Aachen, Germany
Email: {R.Schmitt, Y.Cai}@wzl.rwth-aachen.de

***Abstract*—In processes of the production chain of carbide inserts, such as unloading or packaging, the conformity test of the insert type is performed manually, which causes a statistic increase of errors due to monotony and fatigue of workers as well as the wide variety of insert types. A measuring method is introduced that automatically inspects the chip-former geometry, the most significant quality feature of inserts. The proposed recognition approach is based on local energy model of feature perception and concatenates the phase congruency in terms of local filter orientations into a compact spatial histogram. This method has been tested on several inserts of different types. Test results show that prevalent insert types can be inspected and classified robustly under illumination variations.**

***Index Terms*—optical measurement, industrial image processing, testing and inspection, local energy, phase congruency, feature description, Gabor wavelets, illumination invariance**

I. Introduction

The product spectrum of cutting material manufacturers consists of a large variety of insert geometries, which only differ by small details such as coating colour, edge radius, plate shape and chip-former geometry. An example of a carbide insert and its four remarkable features are shown in Fig. 1. Due to the extensive preparations and the long processing time, different production lots of inserts are combined for an economic and reliable coating. Currently manual conformity test regarding the plate type is used for the unloading of the coating machine. It is cost-intensive and involves a risk of interchanging for the subsequent packaging. Depending on the sorting tasks the error rate of manual insert sorting could be 5 to 35 percent [1] [2]. Generally machining with a false carbide insert type causes dimension faults in mechanical work pieces. For this reason, this series process could be negatively affected and it causes unnecessary costs.

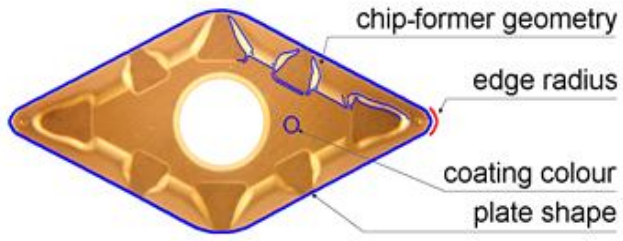

Figure 1. A carbide insert and its four features

The insert quality features, coating colour, edge radius and plate shape, could be inspected easily with simple image processing methods [2] [3]. The chip-former geometry turns out to be the most distinctive feature and a significant quality characteristic of carbide inserts. Compared to the other three quality features the chip-former geometry is the single feature, which represents an insert uniquely. In this paper, a measuring method is proposed for inspection and classification of the chip-former geometry. This technique realises a robust and automated measurement and inspection of carbide insert in the production line.

II. Related Work

A. Object Recognition

Object recognition in computer vision is the task of finding given objects in an image or video sequence and classifying them into the known object types. The central problem of object recognition is how the regularities of objects are extracted and recognized from their images, taken under different lighting conditions and from various perspectives [4] [5] [6]. To capture the object characteristics, early recognition algorithms are focused on using geometric models of objects [7]. While the 3D model of the object being recognized is available, the 2D representation of the structure of an object in image is compared with the 2D projection of the geometric model [7]. Other geometry-based methods try to extract geometric primitives (e.g., lines, circles, etc.) that are invariant to viewpoint change [8]. However, it has been shown that such primitives can only be reliably extracted under limited conditions (controlled variation in lighting and viewpoint with certain occlusion) [7]. In contrast to early geometry-based object recognition works, most recent efforts are concentrated on appearance-based techniques [9]. The underlying idea of this approach is to extract features, as the representations of object, and compare them to the features in the object type database, while 2D representations of objects viewed from different angles and distances are available. Especially the local descriptors [10] [11] [12] do not require object segmentation and describe and represent objects with features directly extracted from local image regions. Local descriptors can be computed efficiently, are resistant to partial occlusion, and are relatively insensitive to changes in viewpoint [13].

B. Local Feature Descriptor

There are two considerations to using local descriptors. First, the interest points are localized in image. The interest points are often occurred at intensity discontinuity and should be stable over transformations [14]. Second, a

DOI: 01.IJEPE.02.03.92_14

description of the image region around the interest point is built. To differentiate interest points reliably from each other, the feature description should ideally be distinctive, concise and invariant over transformations caused by changes in acquisition pose and lighting conditions. While the localization of interest points and feature description are often designed together, the solutions to these two steps are independent [13]. One of the most widely used local feature representation is the scale-invariant feature transform (SIFT) descriptor [11]. The SIFT uses interest points localized at local peaks in scale-space created with a pyramid of difference of Gaussian filters. At each interest point, the local image gradients are sampled and weighted by a Gaussian, and then represented in orientation histograms. The descriptor is formed from a vector containing the values of all the orientation histogram entries. The classification of objects is performed by first matching each interest point in test image independently to the database extracted from template images, and then classifying the objects according to the matches of interest points. Similar to the SIFT descriptor a descriptor is developed by applying principal components analysis (PCA) to the normalized image gradient patch. The dimension of the feature descriptor is more compact than the standard SIFT representation [12]. Another variation of the SIFT method is GLOH (Gradient Location and Orientation Histogram). It considers more spatial regions for the histograms. Through PCA the higher dimensionality of the descriptor is reduced [13].

C. Work Piece Recognition

There are many approaches and methods to resolve the work piece recognition problem [15] [16]. The methods based on local feature descriptors are often applied to 2D images of work pieces. SIFT characteristics are used as matching features in [17]. By using an incline distance as the similarity metrics of image matching, and setting a threshold to delete the false matching points, the matching speed could be improved. A descriptive vector, which consists of object geometric properties, e. g. centroid, orientation angle and height of objects, is produced in [18] to recognize assembly parts. In [19] two image recognition systems based on random local descriptors are introduced. The idea of random descriptors of the image is that each neuron of the associative layer is connected to some randomly selected points on the input image and such neuron its function using the brightness of these points.

III. Machine Vision Prototype

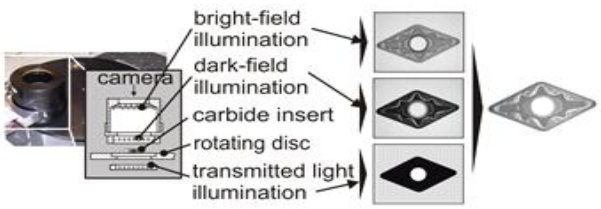

Figure 2. Developed Illumination unit

The developed prototype consists of the following hardware modules: illumination unit, camera/optic-system and mechanical system. To record all necessary information for the measurement and classification task, the illumination concept of the insert inspection system employs a combination of three different lighting types (Fig. 2). The diffuse LED ring light as bright-field illumination provides a homogeneous illumination of the entire insert without disturbing reflections on the surface. By means of an LED ring light as dark-field illumination, a significant highlighting of chip-former geometry is enabled. The use of an LED-based transmitted light illumination ensures a clear outer contour of the insert. To simulate an inline inspection of carbide inserts in the production line, a practically relevant mechanical system is designed and implemented (Fig. 3). The camera/optic-system and the insert handling system are synchronised by a trigger, which is associated with a motorized rotating disc. As the insert comes into the camera view field, the reflexion light sensor releases the image acquisition. As a result, a well-focused insert image is provided.

IV. Classification of Insert Geometry

As shown in Fig. 1, the chip-former geometry appears exactly at the image edges of the insert surface. Based on this fact two methods (template matching with Hausdorff distance [20] and the geometric moments [21]), which classify shapes with features extracted from the image edges, are tested for the chip-former geometry recognition.

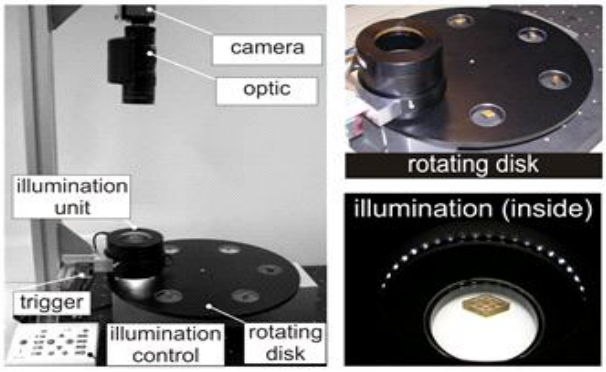

Figure 3. Machine vision prototype for insert inspection

The test result of example insert sets demonstrates that it is not possible to robustly classify the insert sets with extracted chip-former geometries, because these methods normally use gradient-based edge detection operations (such as Canny edge detector [21]) and the detected edges deviate considerably depending on illumination variations and noise [3]. The experiments with local feature descriptors such as SIFT method [11] and PCA-SIFT [12] confirmed that their feature representations are partially dependent on illumination changes [3]. In contrast to the gradient-based methods, phase congruency is a quantity that is invariant to changes in image brightness or contrast [22]. Furthermore, phase congruency is a dimensionless measure, whose value varies from a maximum of 1 (indicating a very significant feature) down to 0 (indicating no significance). This absolute measure of the

DOI: 01.IJEPE.02.03.92_14

significance of feature could be universally applied to any image irrespective of illumination [22]. The application of local feature descriptor which is based on phase congruency of image perception to face recognition shows this method is insensitive to illumination and other common variations in facial appearance [23] [24]. For an illumination-robust classification of chip-former geometry of the carbide insert, an approach is developed based on phase congruency. After the phase congruencies of image are calculated (A. Phase congruence), a feature histogram is formed, which builds a distinctive and concise description of insert image (B. Feature histogram). Based on these extracted feature descriptors, the test insert is identified with an object matching method (C. Matching).

A. Phase Congruence

The local energy model, which has been developed in [25], postulates that significant image features are perceived at points where the Fourier components of the image are maximally in phase. This model defines a phase congruency function in terms of the Fourier series expansion of a signal at location x as

$$PC(x) = \frac{\sum_n A_n(x)\cos(\phi_n(x) - \bar{\phi}(x))}{\sum_n A_n(x)}, \quad (1)$$

where A_n and ϕ_n represent respectively the amplitude and local phase of the n-th Fourier component at position x. $\bar{\phi}(x)$ is the amplitude weighted mean local phase angle of all the Fourier terms at the point

$$\bar{\phi}(x) = \frac{\sum_n A_n(x)\phi_n(x)}{\sum_n A_n(x)}. \quad (2)$$

The local energy function is defined for signal $I(x)$ as

$$E(x) = \sqrt{F^2(x) + H^2(x)} \quad (3)$$

where $F(x)$ is the original signal $I(x)$ with its direct component removed and $H(x)$ is the Hilbert transform of $F(x)$ [26]. As shown in Fig. 4, local energy is equal to phase congruency scaled by the sum of the Fourier amplitudes

$$E(x) = PC(x)\sum_n A_n(x) \quad (4)$$

Thus, the local energy is directly proportional to the phase congruency and energy peaks are corresponding to peaks in phase congruency, which indicate image feature significance. In (1) and (4) the phase congruency is defined as the ratio of $E(x)$ to the overall path length of local Fourier components (Fig. 4). Thus, the degree of phase congruency is independent on the overall magnitude of the signal. This provides invariance to image illumination variations [22]. The phase congruency in (1) is sensitive to noise and results in poor localization [23]. To compensate these effects, this measure is extended by consisting of sine of the phase deviation, including a proper weighing of the frequency spread $W(x)$ and also a noise cancellation factor T. This modified measure, as given in (5), is more sensitive and yields good localization of blurred features [22]

$$PC(x) = \frac{\sum_n W(x)\lfloor A_n(x)\Delta\Phi_n(x) - T\rfloor}{\sum_n A_n(x) + \varepsilon} \quad (5)$$

$$\Delta\Phi_n(x) = \cos(\phi_n(x) - \bar{\phi}(x)) - \left|\sin(\phi_n(x) - \bar{\phi}(x))\right| \quad (6)$$

(1) is expanded by

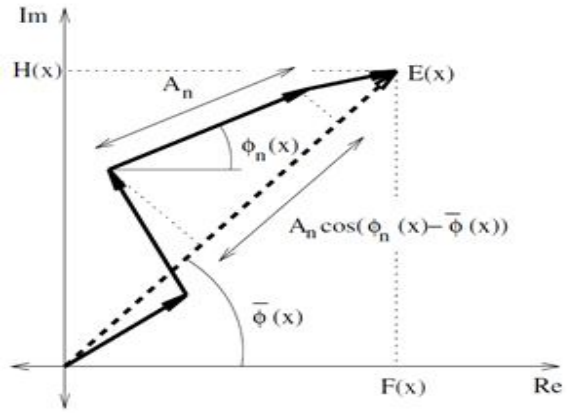

Figure 4. Local energy and phase congruency from signal Fourier components

To obtain frequency information (magnitude A_n and phase ϕ_n) local to a point in a signal, logarithmic Gabor wavelet filters are used [27]. This linear phase filters are in symmetric/ antisymmetric quadrature pairs and preserve phase information. Compared to Gabor filters, the logarithmic filters maintain a zero direct component in the even-symmetric filter for arbitrarily large bandwidth filters [22]. This one-dimensional symmetric/antisymmetric Gabor filter pairs are extended into two dimensions by simply applying Gaussian as spreading function across the filter perpendicular to its orientation. By using a bank of these two-dimensional filters, which is designed to tile the frequency plane uniformly, features at all orientations can be detected. The energies over all orientations are summed and normalized by dividing by the sum over all orientations and scales of the amplitudes of the individual wavelet responses at that location x in the image. In this way the two-dimensional phase congruency is calculated as

$$PC_2(x) = \frac{\sum_o \sum_n W_o(x)\lfloor A_{n,o}(x)\Delta\Phi_{n,o}(x) - T_o\rfloor}{\sum_o \sum_n A_{n,o}(x) + \varepsilon} \quad (7)$$

where o denotes the filter orientation.

B. Feature Histogram

Based on the local energy model, a histogram is generated by accumulating the calculated phase congruency (Fig. 5,

DOI: 01.IJEPE.02.03.92_14

middle) along each filter orientation on different regions of the image (image partition in our experiments is shown in Fig.6). Before histogram conformation, an orientation label map for the image is obtained based on local energies. To each pixel a consistent orientation o is assigned, at which it has largest energy across all scales

$$L(x) = \arg\max_{o}\left(W_o(x)\sum_n \lfloor A_{n,o}(x)\Delta\Phi_{n,o}(x) - T_o \rfloor\right) \quad (8)$$

For a single image region, a local histogram is created by summarizing the phase congruencies in certain orientation. These local histograms extracted from different regions of the image are then concatenated together, while keeping the spatial relationship between image parts. This produces the histogram for the whole image (Fig. 5, right)

$$H_{r,o} = \sum_x \omega_r(x) PC_2(x)\delta(L(x) - o) \quad (9)$$

where

$$\delta(L(x) - o) = \begin{cases} 1 & \text{for L(x) = o} \\ 0 & \text{for L(x) } \neq \text{ o} \end{cases} \quad (10)$$

and $\omega_r(x)$ is Gaussian weighing function centred at region r. This weight provides soft margins across

Figure 5. Phase congruency-based feature descriptor

histogram bins by small weighted overlap among neighbouring regions to overcome the problems induced by scale variations [23].

C. Matching

The classification of chip-former geometry is performed by matching the histogram of test image independently to the database extracted from template images. For this matching the Jeffrey divergence between corresponding feature vectors is calculated, which is used as measure of similarity [28]

$$d(H,K) = \sum_r^R \eta_r \sum_o^O \left(h_{r,o}\log\frac{h_{r,o}}{m_{r,o}} + k_{r,o}\log\frac{k_{r,o}}{m_{r,o}}\right) \quad (11)$$

where $h_{r,o}$ is the corresponding bin in histogram H for region r and orientation o, $m_{r,o}$ is the mean of bins and given by

$$m_{r,o} = \frac{(h_{r,o} + k_{r,o})}{2} \quad (12)$$

and $\eta_r \in [0,1]$ is weight of each region and assigns to each region different influence on the divergence calculation (Fig. 6).

V. Experimental Results

In this section, we present the experimental results of the insert inspection, examine the classification rate and also evaluate the feature invariance to image transformations.

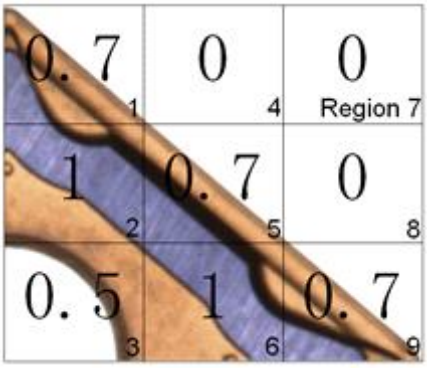

Figure 6. Image partition and weights of regions for divergence calculation

Figure 7. A test set of carbide inserts

A. Test Sets

We evaluate the inspection method on real carbide inserts, which have different geometries from different manufacturers. An example test set is given in Fig. 7, which contains 14 insert types. For each type there are three insert examples available. We use two examples of each type (namely two test insert sets) to acquire test insert images. For each insert we get with two different scales, under three varied illumination conditions and at two random positions 2x3x2=12 test images. Corresponding to the insert set and the image scale the test images are grouped in 2x2=4 test image sets.

With these insert sets a test for the classification of the chip-former geometry can be estimated from a sample of 14x2x12=336 test images of real carbide inserts. By using the remaining insert set we acquire one template for each insert type, which is grabbed under a controlled illumination condition.

B. Test Results

In our experiments, local histograms with 4 bins corresponding to 4 filter orientations on 9 image regions are extracted, which create a 36-dimensional feature vector. The phase congruency is achieved by convolving the image with a bank of Gabor wavelets kernels tuned to 4 frequency scales and 6 orientations. In Fig. 8 the test results of the example test set of carbide inserts in Fig. 7 are presented. Our experiments show, that the descriptor extracted from the phase congruency is relatively robust in the production environment and achieves an overall average classification rate of 97.02%.

DOI: 01.IJEPE.02.03.92_14

It is similarly effective compared to the method developed in [3], which uses the SIFT local feature descriptor [11], diminishes the false matched feature vectors with a location consistency check of feature points and provides a classification rate of 98.56% for the test on the same image sets.

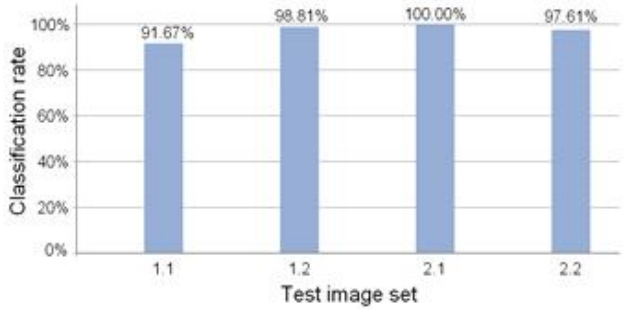

Figure 8. Classification rates of four test image sets

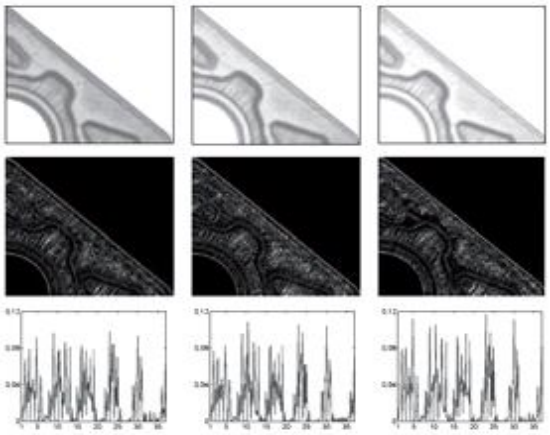

Figure 9. Comparison of phase congruencies under different illuminations

C. Feature Invariance

In this section the robustness of the feature descriptor with respect to image translation, rotation, scale change and illumination change is evaluated. In Fig. 9 the phase congruencies for insert images acquired under three illumination settings are shown and the invariance to image illumination can be indicated exemplarily. The phase congruencies show no obvious differences due to the changed lighting situations and the histograms are also very similar. Because the histogram is concatenated after the spatial relationship of image regions, this phase congruency-based feature descriptor is not invariant to translation and rotation. Our experiments show that the classification rate drops strongly from a rotation of 5° and from a translation of a tenth of the image size. In order to analyse the scale invariance, test images are compared with differently scaled reference data sets. The corresponding classification rates are shown in Fig. 10. The blue curves present the results from the matching with the low-resolution reference (horizontal resolution 310 pixels), while red curves illustrate the rates from the comparison with the high-resolution reference (horizontal resolution 435 pixels). High-resolution image contains more details. These subtleties prepare again image features, which are not represented in low-resolution image. The developed descriptor is sensitive to these subtleties and therefore the invariance to image scale is expected only in a closely limited range.

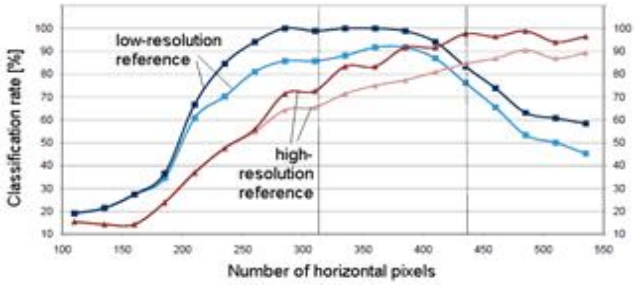

Figure 10. Classification rates for comparison with differently scaled references

Conclusions

This paper describes a measuring method for an inspection in carbide insert production. The experiments on real inserts show that the proposed, on phase congruency-based method can achieve both accuracy for chip-former geometry recognition and high robustness to illumination.

References

[1] T. Pfeifer, *Production Metrology*, Oldenbourg, 2002.
[2] R. Hermes, *Entwicklung flexibler Bildverarbeitungsketten zur Klassifikation und Verschleißmessung an Wendeschneidplatten*, PhD thesis, RWTH University, 2007 (in German).
[3] R. Schmitt, Y. Cai, and T. Aach, "A Priori Knowledge-Based Recognition and Inspection in Carbide Insert Production", *Proceedings IEEE International Conference on Computational Intelligence for Measurement Systems and Applications* (CIMSA 2010), pp. 339-342, 2010.
[4] R. O. Duda, P. E. Hart and D.G. Stork, *Pattern Classification*, 2nd edition, John Wiley & Sons, 2001.
[5] C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006.
[6] R. Brunelli, *Template Matching Techniques in Computer Vision: Theory and Practice*, Wiley, 2009.
[7] J. Mundy, "Object Recognition in the Geometric Era: a Retrospective", in *Toward category-level object recognition*, edited by J. Ponce, M. Hebert, C. Schmid and A. Zisserman, Springer, pp. 3-29, 2006.
[8] D. Zhang, G. Lu, "Review of shape representation and description techniques", *Pattern Recognition*, Vol. 37, No. 1, pp. 1-19, 2004.
[9] H. Murase and S. K. Nayar, "Visual learning and recognition of 3-D objects from appearance", *International Journal of Computer Vision*, 14, pp. 5-24, 1995.
[10] C. Harris and M. Stephens, "A combined corner and edge detector". In *Alvey Vision Conference*, pp. 147–151, 1988.
[11] D. G. Lowe, "Distinctive image features from scale-invariant keypoints". *International Journal of Computer Vision*, pp. 91-110, 2004.
[12] Y. Ke and R. Sukthankar, "PCA-SIFT: A More Distinctive Representation for Local Image Descriptors," *Proc. Conf. Computer Vision and Pattern Recognition*, pp. 511-517, 2004.

DOI: 01.IJEPE.02.03.92_14

[13] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(10), pp. 1615-1630, 2005.
[14] K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaffalitzky, T. Kadir, and L. Van Gool, "A comparison of affine region detectors". *International Journal of Computer Vision*, 65(1/2), pp. 43-72, 2006.
[15] G. K. Toledo, E. Kussul, T. Baidyk, "Neural classifier for micro work piece recognition", *Image and Vision Computing*, 24, pp. 827-836, 2006.
[16] A. Shafi, "Machine vision in automotive manufacturing", *Sensor Review*, Vol. 24, Iss. 4, pp. 337-342, 2204.
[17] Y. Wang, W. Fu, H. Zhu and Y. Liang, „A Method for Recognition of Work-pieces Based on Improved SIFT Characteristics Matching", *Proceedings of the 2009 WRI World Congress on Software Engineering*, pp. 29-33, 2009.
[18] M. Peña, I. López, R. Osorio, "Invariant Object Recognition Robot Vision System for Assembly", *Proceedings of the Electronics, Robotics and Automotive Mechanics Conference*, pp. 30-06, 2006.
[19] E. Kussul, T. Baidyk, D. Wunsch, O. Makeyev and A. Martín, "Image Recognition Systems Based on Random Local Descriptors", *Proceedings of Joint Conference on Neural Networks*, pp. 2415-2420, 2006.
[20] A. Hornberg, *Handbook of Machine Vision*, Wiley-VCH, 2006.
[21] B. Jähne, *Digital Image Processing*, Springer, 2002.
[22] P. D. Kovesi, "Image Features from Phase Congruency", *Videre: Journal of Computer Vision Research*, MIT Press. Vo. 1, Nr. 3, pp. 1-26, 1999.
[23] S. Sarfraz and O. Hellwich, "Head Pose Estimation in Face recognition Across Pose Scenarios", *Proceedings of VISAPP 2008, Int. conference on Computer Vision Theory and Applications*, pp. 235-242, 2008.
[24] K. Anusudha1, S. A. Koshie, S. S. Ganesh and K. Mohanaprasad, "Image Splicing Detection involving Moment-based Feature Extraction and Classification using Artificial Neural Networks", ACEEE Int. J. on Signal & Image Processing, Vol. 01, No. 03, pp. 9-13, 2010.
[25] M. C. Morrone and R. A. Owens, "Feature Detection from Local Energy", *Pattern Recognition Letters*, 6, pp. 303-313, 1987.
[26] S. Venkatesh and R. A. Owens, "An Energy Feature Detection Scheme", *International Conference on Image Processing*, pp. 553-557, 1989.
[27] K. Gröchenig, *Foundations of Time-frequency Analysis*, Birkhäuser, 2001.
[28] Y. Rubner, C. Tomasi and L. J. Guibas, "The earth Mover's Distance as a Metric for Image Retrieval", *International Journal of Computer Vision*, 40(2), pp. 99–121, 2000.

DOI: 01.IJEPE.02.03.92_14